UNIVERSITY OF NOVI SAD
TECHNICAL FACULTY
"MIHAJLO PUPIN"
ZRENJANIN

ITROCONFERENCE 9.0
INFORMATION TECHNOLOGY AND EDUCATION DEVELOPMENT

PROCEEDINGS

ZRENJANIN, June 2018

UNIVERSITY OF NOVI SAD
TECHNICAL FACULTY "MIHAJLO PUPIN"
ZRENJANIN
REPUBLIC OF SERBIA

IX INTERNATIONAL CONFERENCE OF

INFORMATION TECHNOLOGY AND DEVELOPMENT OF EDUCATION ITRO 2018

PROCEEDINGS OF PAPERS

IX MEĐUNARODNA KONFERENCIJA

INFORMACIONE TEHNOLOGIJE I RAZVOJ OBRAZOVANJA ITRO 2018

ZBORNIK RADOVA

ZRENJANIN, JUNE 2018

Publisher and Organiser of the Conference:
University of Novi Sad, Technical faculty „Mihajlo Pupin", Zrenjanin, Republic of Serbia

For publisher:
Dragica Radosav, Ph. D, Professor, Dean of the Technical faculty „Mihajlo Pupin", Zrenjanin, Republic of Serbia

Editor in Cheaf - President of OC ITRO 2018:
Vesna Makitan, Ph. D, Assistant Professor,

Proceedings editor:
Marjana Pardanjac, Ph. D, Professor

Technical design:
Ivan Tasic, Ph. D, Professor;
Dusanka Milanov MSc, Assistant

Lecturer:
Erika Tobolka, Ph. D, Professor

Circulation: **50**
ISBN: 978-86-7672-310-2

By the resolution no. 142-451-607/2018-01/02, Autonomous Province of Vojvodina, Provincial Secretariat For Science and Technological Development donated financial means for printing this Conference Proceedings.

The Conference is supported by the Autonomous Province of Vojvodina

CIP - Каталогизација у публикацији
Библиотека Матице српске, Нови Сад

37.01:004(082)
37.02(082)

INTERNATIONAL Conference on Information Technology and Development of Education ITRO (9 ; 2018 ; Zrenjanin)
Proceedings of papers [Elektronski izvor] / IX International Conference on Information Technology and Development of Education ITRO 2018 = Zbornik radova / IX međunarodna konferencija Informacione tehnologije i razvoj obrazovanja ITRO 2018, Zrenjanin, June 2018. - Zrenjanin : Technical Faculty "Mihajlo Pupin", 2018. - 1 elektronski optički disk (CD-ROM) : ilustr. ; 12 cm

Nasl. sa naslovnog ekrana. - Tiraž 50. - Bibliografija uz svaki rad.

ISBN 978-86-7672-310-2
I. Međunarodna konferencija Informacione tehnologije i razvoj obrazovanja ITRO (9 ; 2018 ; Zrenjanin) International Conference on Information Technology and Development of Education ITRO (9 ; 2018 ; Zrenjanin)
a) Информационе технологије - Образовање - Зборници b) Образовна технологија - Зборници
COBISS.SR-ID 324298503

PARTNERS INTERNATIONAL CONFERENCE

South-West University „Neofit Rilski”
Faculty of Education, Blagoevgrad,
Republic of Bulgaria

SOUTH WEST UNIVERSITY
"NEOFIT RILSKI"

Faculty of Electrical Engineering and Informatics
Department of Computers and Informatics of Kosice
Slovak Republic

TECHNICKÁ UNIVERZITA
FEI
KOŠICE

University Goce Delcev Stip
Republic of Macedonia

УНИВЕРЗИТЕТ
„ГОЦЕ ДЕЛЧЕВ“
ШТИП

THE SCIENCE COMMITTEE:

Dragoslav Herceg, Ph.D, Professor, Faculty of Science, Novi Sad, Serbia – Chairman

Marina Čičin Šain, Ph.D, Professor, University of Rijeka, Croatia

Anton Vukelić, Ph.D, Professor, Faculty of Philosophy, Croatia

Ion Dzitac, Ph.D, Professor, Department of Mathematics-Informatics, Aurel Vlaicu University of Arad, Romania

Sashko Plachkov, Ph.D, Professor, South-West University "Neofit Rilski" /Department of Education, Blagoevgrad, Republic of Bulgaria

Sulejman Meta, Ph.D, Professor, Faculty of Applied Sciences, Tetovo, Macedonia

Márta Takács, Ph.D, Professor, Óbuda University, John von Neumann Faculty of Informatics, Budapest, Hungary

Nina Bijedić, Ph.D, Professor, Applied mathematics, Bosnia and Herzegovina

Viorel Negru, Ph.D, Professor, Department of Computer Science, West University, Timisoara, Romania

Mirjana Segedinac, Ph.D, Professor, Faculty of Science, Novi Sad, Serbia

Milka Oljača, Ph.D, Professor, Faculty of Philosophy, Novi Sad, Serbia

Dušan Starčević, Ph.D, Professor, Faculty of Organizational Sciences, Belgrade, Serbia

Dobrivoje Mihajlović, Ph.D, Professor, Faculty of Organizational Sciences, Belgrade, Serbia

Kosta Voskresenski, Ph.D, Professor, Technical Faculty "Mihajlo Pupin" Zrenjanin, Serbia

Josip Ivanović, PhD, Professor, Hungarian Language Teacher Training Faculty, Subotica, Serbia

Ivanka Georgieva, Ph.D, South-West University "Neofit Rilski", Faculty of Engineering, Blagoevgrad, Republic of Bulgaria

Miodrag Ivković, Ph.D, Professor, Technical Faculty "Mihajlo Pupin" Zrenjanin, Serbia

Momčilo Bjelica, Ph.D, Professor, Technical Faculty "Mihajlo Pupin" Zrenjanin, Serbia

Dragica Radosav, Ph.D, Professor, Technical Faculty "Mihajlo Pupin" Zrenjanin, Serbia

Dragana Glušac, Ph.D, Professor, Technical Faculty "Mihajlo Pupin" Zrenjanin, Serbia

Dijana Karuović, Ph.D, Professor, Technical Faculty "Mihajlo Pupin" Zrenjanin, Serbia

Ivan Tasić, Ph.D, Professor, Technical Faculty "Mihajlo Pupin" Zrenjanin, Serbia

Vesna Makitan, Ph.D, Assistant Professor, Technical Faculty "Mihajlo Pupin" Zrenjanin, Serbia

Marjana Pardanjac, Ph.D, Professor, Technical Faculty "Mihajlo Pupin" Zrenjanin, Serbia

Snežana Babić Kekez, Ph.D, Professor,Technical Faculty "Mihajlo Pupin" Zrenjanin, Serbia

Erika Tobolka, Ph.D, Professor, Technical Faculty "Mihajlo Pupin" Zrenjanin, Serbia

Stojanov Željko, Ph.D, Professor, Technical Faculty "Mihajlo Pupin" Zrenjanin, Serbia

Brtka Vladimir, Ph.D, Professor, Technical Faculty "Mihajlo Pupin" Zrenjanin, Serbia

Kazi Ljubica, Ph.D, Assistant Professor, Technical Faculty "Mihajlo Pupin" Zrenjanin, Serbia

Berković Ivana, Ph.D, Professor, Technical Faculty "Mihajlo Pupin" Zrenjanin, Serbia

Nikolić Milan, Ph.D, Professor, Technical Faculty "Mihajlo Pupin" Zrenjanin, Serbia

Stojanov Jelena, Ph.D, Assistant Professor, Technical Faculty "Mihajlo Pupin" Zrenjanin, Serbia

THE ORGANIZING COMMITTEE:

Vesna Makitan, Ph.D, Ass. Professor, Technical Faculty "Mihajlo Pupin" Zrenjanin, R. of Serbia - Chairman of the Conference ITRO 2018
Dragica Radosav, Ph.D, Professor, Technical Faculty "Mihajlo Pupin" Zrenjanin, R. of Serbia
Dijana Karuovic, Ph.D, Professor, Technical Faculty "Mihajlo Pupin" Zrenjanin, R. of Serbia
Marjana Pardanjac, Ph.D, Professor, Technical Faculty "Mihajlo Pupin" Zrenjanin, R. of Serbia
Ivan Tasic, Ph.D, Professor, Technical Faculty "Mihajlo Pupin" Zrenjanin, R. of Serbia
Dragana Glusac, Ph.D, Professor, Technical Faculty "Mihajlo Pupin" Zrenjanin, R. of Serbia
Erika Tobolka, Ph.D, Professor, Technical Faculty "Mihajlo Pupin" Zrenjanin, R. of Serbia
Dusanka Milanov, MSc, Assistant, Technical Faculty "Mihajlo Pupin" Zrenjanin, R. of Serbia
Vladimir Karuović, Technical Faculty "Mihajlo Pupin" Zrenjanin, R. of Serbia

With this publication, the CD with all papers from the International Conference on Information Technology and Development of Education, ITRO 2018 is also published.

INTRODUCTION

Technical Faculty "Mihajlo Pupin" organized, now the traditional, IX International Conference on Information Technology and Education Development (ITRO 2018), which was held on June 29, 2018.

This year we managed to gather our colleagues, scientists, researchers and students from 10 countries (Serbia, Macedonia, Bulgaria, Bosnia and Herzegovina, Romania, USA, Great Britain, Albania, Montenegro, Slovakia). Many of them have been participating in the work of the Conference for many years and practically they are making an ITRO family. With their papers they managed to present and promote the results of research and scientific work in the field of information technology in education. More than 40 papers have been collected, which will be published in the Proceedings of the Conference website too (http://www.tfzr.rs/itro/index.html).

The main course in the work of the Conference was set up with introductory lectures in which the significance of following topics could be seen:

- Education for modern business and education from the perspective of employers nowadays when every company is directly or indirectly IT company – lecture with the topic "Digital transformation of the society – the role of education" was held by Goran Đorđević, director of the company Consulteer;
- Scientific research work in the field of information technology in education, whose results were published in one of the world's leading magazines – this novelty at the ITRO Conference was introduced by PhD Dragana Glušac with a lecture on "School without walls";
- The latest forms of education and practice of IT experts in the country and abroad – a lecture on the topic "Finding a space for "making" and digital fabrication in the education of Serbia" was held by PhD Dalibor Dobrilović.

The other presented papers have cast light on various aspects of contemporary education in our country and abroad, as well as on the experiences, problems, questions, etc. which are related to them.

The conference was an opportunity to connect again with researchers and scientists from other institutions and countries and ask questions about new forms of cooperation and projects that are relevant to all of us.

The conference was held thanks to the sponsorship of the Provincial Secretariat for Higher Education and Scientific Research, which also traditionally supports ITRO, as well as the Faculty, which provided the necessary technical conditions.

We thank everyone for participating and creating the ITRO tradition.

See you at the next ITRO Conference,

Chairman of the Organizing Committee

PhD Vesna Makitan

We are very grateful to:

Autonomous Province of Vojvodina

for donated financial means which supported printing of the Conference Proceedings and organizing of the Conference.

CONTENTS

INIVITED LECTURE

SCIENTIFIC PAPERS

Classification with ID3 and SMO using Weka

A. Velinov and C. Martinovska Bande
Goce Delcev University, Faculty of Computer Science, Stip, Republic of Macedonia

Abstract – In this paper, we look at ID3 and SMO (SVM) classification algorithms. We used the wine quality dataset that is publicly available. In our research, we classify red wine instances. Data refer to "Vinho Verde", a product that is produced in Minho, a district in northwestern Portugal. According to the obtained results, SMO has more correctly classified instances than ID3, but ID3 has fewer incorrectly classified instances.

I. INTRODUCTION

The number of data is constantly increasing. Data mining refers to extracting or mining the knowledge from large amount of data. The term data mining is appropriately named as 'Knowledge mining from data' or "Knowledge mining" [1]. Knowledge Discovery in Databases is the non-trivial process of identifying valid, novel, potentially useful, and ultimately understandable patterns in data [2]. Figure 1 shows the process for discovering knowledge from databases. First step is to select the data to be used. From databases we can select the tables with records that we will use to extract knowledge. Often tables contain data that does not match our research goals. It is therefore necessary to initially preprocess the data. The transformation phase follows. At this stage, the data is transformed into a form that is suitable for research. The next process is the discovery of patterns from the transformed data. Basically this is the data mining phase. The last step of this process is the evaluation of the knowledge gained from data.

As we have previously stated, the basic purpose of data analysts is to extract data schemes. Usually the data are stored electronically, and the search is done with the help of a computer. Patterns in data can be automatically identified, validated and used for future predictions. Analysts are trying to increase opportunities for finding patterns in the data. If this process is automated, it will ease the work and in the future we can more easily and more efficiently extract knowledge from the data. According to some analyzes, the amount of data stored in databases doubles every 20 months [2]. This is a big step in increasing the data. Although the number of data increases, with the discovery of new search engines, the opportunities for gaining knowledge from data are increasing. If the data are analyzed intelligently, valuable resources can be

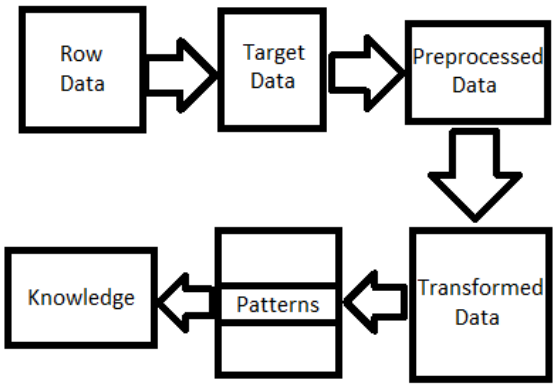

Figure 1. Process for discovering knowledge in databases

obtained, which can later be used for some future plans.

Data mining is about solving problems by analyzing data that are already present in databases [3]. Discovering patterns in data is one of the characteristics of data mining.

II. CLASSIFICATION

Classification is one of the Data Mining techniques that is mainly used to analyze a given dataset, takes each instance of it and assigns this instance to a particular class [4]. This process aims to make the classification error very small. The classification allows the extraction of models that define the classes for a given data set.

There are three different learning approaches that are used for data mining: supervised learning and unsupervised learning [5]. The algorithms for supervised learning use set of examples that have known labels. If the labels are nominal value, then it is a classification task. If the labels are numerical values, then it is regression task. Basically, the classification and regression are two types of supervised learning. When we have a classification problem, then the output variable is a category such as "green", "blue", "disease", "no disease". In regression, as output we have real values. Examples are "dollars", "weight" and so on. This type of learning uses input variables (x)

and output variables (y). An algorithm is used to learn the mapping function from the input to the output.

$$y=f(x) \quad (1)$$

The idea is that if we have new input data (x) in the future, we can predict the outputs from this data (y). This learning algorithm looks a lot like the supervision of learning by the teacher. That is why this is called supervised learning. The correct answers are familiar to us. The algorithm uses training data to make predictions. This process is iterative. The algorithm is corrected by the teacher. When it reaches an acceptable level of performance, it stops. Supervised learning is the most used practical machine learning.

The labels of the examples in the data sets in unsupervised learning are unknown. This algorithm uses the similarity of attribute values to group the examples. If there are only input values (x) and no corresponding output variables, then it is unsupervised learning. This type of learning, tend to model the distribution in the data in order to learn more about data. Unlike supervised learning, there is no correct answers and there is no teacher. Algorithms use own sources to discover and to present the interesting structure in the data. Examples of unsupervised learning are: clustering and association. Clustering algorithms are used to discover the inherent groupings in the data. An example of this is grouping customers by purchasing behavior. Association tend to discover the rules that describe large portions of data. An example of association is: "People that buy fish also tend to buy salt". Some of the more familiar algorithms for unsupervised learning are: "k-means" for clustering problems and "Apriori" for association rule learning problems [6]. Semi-supervised learning has a small subset of labelled examples. It also has a lot of unlabeled examples. These problems are found between supervised learning and unsupervised learning. An example for this is when we have a folder with documents that are labeled and the majority are unlabeled. Here, unsupervised learning is used to discover and to learn the structure in the input variables. Supervised learning is used to make best predictions for the unlabeled data.

In this paper, we will look at the classification as well as the algorithms for classification. As we have said before, classification uses given input to predict a certain outcome. The algorithm processes

TABLE I. TRAINING AND PREDICTION SET

Fixedacidity	7.4	11.2	7.3	7.6
Volatileacidity	0.7	0.28	0.65	0.3
Citricacid	0	0.56	0	0
Residualsugar	1.9	1.9	1.2	1.4
Chlorides	0.076	0.075	0.065	0.07
Freesulfurdioxide	11	17	15	10
Totalsulfurdioxide	34	60	21	34
Density	0.9978	0.998	0.994 6	0.88
pH	3.51	3.16	3.39	3.2
Sulphates	0.56	0.58	0.47	0.5
Alcohol	9.4	9.8	10	9.7
Quality	5	6	7	?

a training set in order to predict the outcome. The training set contains a set of attributes and the respective outcome, usually called goal or prediction attribute [7]. The relationship between the attributes is used to predict the outcome. Next, this algorithm uses prediction set. It contains the same set of attributes except the prediction attribute (it is not yet known). The algorithm produces predictions based on the analysis of the input. Using the accuracy of the algorithm we know if it is good or not.

In the Table 1, we can see the training set and prediction set, which contains data about wine quality. Here the prediction attribute (quality) indicates the quality of wine. The prediction set is given in the last column of the table. The goal is to use the data from the training set, to determine the wine quality of the example from prediction set.

Classification often uses prediction rules to

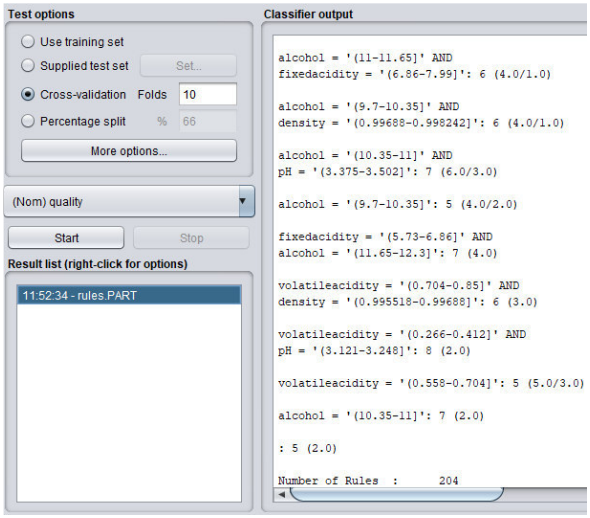

Figure 2. Rules for the wine quality dataset

express knowledge [7]. Prediction rules consist of IF-THEN statements. IF expression consists of conjunction of conditions. After IF part follows THEN which contains the prediction. Some of the rules for the wine quality dataset that are obtained using the PART algorithm in Weka are given in Figure 2. There are 204 rules in total.

Often the prediction rules are very large. How well predictions are done, is measured in percentage of predictions hit against the total number of predictions.

There are several classification algorithms: Decision Trees, K-Nearest Neighbor, Support Vector Machines, Naïve Bayesian Classification, Neural Networks. In this paper, we considered the decision trees and Support Vector Machines.

III. DECISION TREES

The decision tree is one of the most commonly used classification algorithms. Decision tree is hierarchical tree structure that is used to classify classes based on series of questions (or rules) about the attributes of the class [8]. The attributes can be nominal, ordinal, binary and quantitative values. The classes must be qualitative types. The decision tree produces sequence of rules based on data of attributes and its classes. Basically the sequence of rules are series of questions. These questions are used to recognize the class. We ask questions and we answer until we make a conclusion about the class to which the record belongs. These questions can be presented as form of decision tree. It is hierarchical structure which consist of nodes and directed edges. There are three types of nodes: root node, internal nodes and leaf or terminal nodes. Root node is the top node in the three. It has no incoming edges and zero or more outgoing edges. Internal nodes have exactly one incoming edge and two or more outgoing edges. Terminal nodes have one incoming edge and no outgoing edges. In decision trees, each terminal node is a class (class label). Internal nodes represent attribute test conditions, to separate instances with different characteristics. Once the decision tree is constructed, it is easy to classify a test record. First, we start from the root node and we apply the test condition to the record (We ask the questions). Based on the outcome of the test, we follow the appropriate branch. After that, we have another internal node (another test condition) or terminal (leaf) node. If the last step is terminal node, then the class label is assigned to the record. In our paper, for decision three, we look at the ID3 algorithm.

A. ID3 algorithm

ID3 is a simple decision tree learning algorithm developed by Ross Quinlan (1983) [9]. This algorithm uses the greedy search technique on a given dataset, to test each attribute at every tree node. ID3 uses a metric known as information gain to select the attribute that is most useful for classifying a given dataset. We used this metric to minimize the depth of the tree (minimize asked questions). It is a function that measure which questions provide the most balanced splitting [9]. The information gain measures how well a given attribute separates training examples into class labels [10]. We selected the attribute with highest information gain (most useful information for classification). In order to define the gain, we firstly introduce the notion of entropy from information theory. The entropy measures the amount of information in an attribute. On a given set S, which contains positive and negative instances, the entropy is [10]:

$$\text{Entropy}(S) = \sum -p(I) \log 2\ p(I) \qquad (2)$$

or

$$\text{Entropy}(S) = - P(\text{positive}) \log 2 P(\text{positive}) - P(\text{negative}) \log 2 P(\text{negative})$$

For example, if S is a collection of 15 examples with 10 YES and 5 NO examples then [10]:

$$\text{Entropy}(S) = - (10/15)\ \text{Log2}\ (10/15) - (5/15)\ \text{Log2}\ (5/15)$$

If all members of S belong to the same class, then the entropy is 0. It is a case when we have perfectly classified instances. If the entropy is 1, then the classification is totally random.

The information gain is expected reduction of entropy related to specific attribute when we split a decision tree node [9]. The information gain Gain(S,A) of example set S, on attribute A is [10]:

$$\text{Gain}(S, A) = \text{Entropy}(S) - \sum ((|Sv| / |S|) * \text{Entropy}(Sv)) \qquad (3)$$

We calculate the information gain for each attribute and we use the highest gain as the decision node. The attribute that has a highest gain will be root node. We use the gain to branch the decision tree.

To evaluate the precision and recall of the classifiers we must compute several measures. Precision of the classifier can be interpreted as

probability of a instance classified in the current class actually to belong to that class and is defined as:

$$P = \frac{TP}{TP + FP} \quad (4)$$

where FP (False Positive) is the number of instances incorrectly labeled as belonging to the current class and TP (True Positive) is the number of correctly labeled instances that belong to the current class. The sensitivity or recall of the classifier denotes the probability that a instance of a current class is correctly classified. This measure is computed according to the formula:

$$R = \frac{TP}{TP + FN} \quad (5)$$

FN (False Negative) is the number of instances that belong to the current class incorrectly labeled as belonging to other classes, and TP has the same meaning as defined in (4).

If we combine previous two metrics, we can estimate the efficiency of the classifier. For that purpose, measure F1 is computed as harmonic mean of precision and recall. F1 is calculated according to the following formula:

$$F1 = \frac{2RP}{R + P} = \frac{2 \times TP}{2 \times TP + FP + FN} \quad (6)$$

IV. SUPPORT VECTOR MACHINES

Support Vector Machines (SVM) is a classifier defined by a separating hyperplane [11]. SVM represents the separating hyperplane (decision boundary) using a subset of the training examples. This technique has been applied to a number of practical application and it has roots in statistical learning theory. Using training data, this algorithm can in the future determine a hyperplane that classifies new instances. This hyperplane in a two dimensional space is a line that divide the plane in two parts (two classes). We can use SVM for classification also with high-dimensional data (solving the problem with dimensionality). In Figure 2 we can see a classification problem consisting of two classes. One of the classes is set of circles and another is set of squares. These two classes are separated by line. All objects on the left side of the line belong to the class squares. All objects on the right side of the red line belong to the class circles. This separation of classes is a

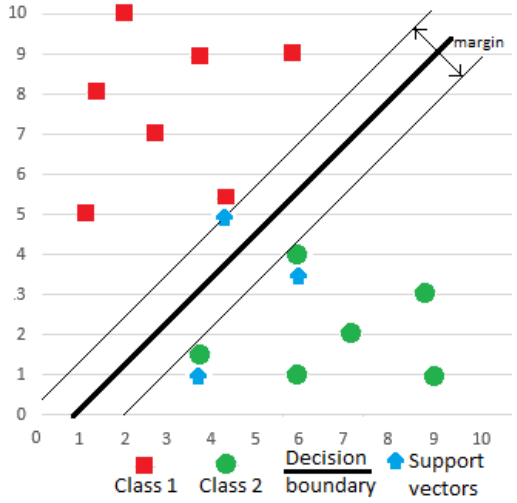

Figure 3. SVM. Separation of classes

SVM feature. SVM finds a line (hyperplane) that separate the classes. The goal of SVM is to find the best classification function to separate the training examples into two classes [5]. This classification function can be determined geometrically. This function corresponds to a line in two dimensional space and it passes through the middle of the two classes (black line of Figure 2). Using this function in the future, new data instances can be classified. If we look at the example in Figure 3, we can see that other lines can also be drawn that also separate objects into two classes. SVM classification function maximize the margin between the two classes. The margin is a space between the two classes that is separated by a line. This is a concept of maximum margin hyperplanes. SVM enables the creation of classifiers that maximize the margin of hyperplanes to minimize the worst-case generalization errors. An example of such a classifier is the linear SVM. It searches for a hyperplane with the largest margin. This classifier is also known as maximal margin classifier. The decision boundary of a linear classifier can be presented as:

$$w*x+b=0, \quad (7)$$

Here, w and b are parameters of the linear decision boundary model.

Margin maximization is equivalent to the minimization of the following objective function:

$$f(w) - \frac{\|w\|^2}{2} \quad (8)$$

The learning task in SVM (Linear separable case) can be formalized as:

$$\min_{w} \frac{||w||^2}{2}$$
$$\text{subject to } y_i(w*x_i+b)\geq 1,\ i=1,2,\ldots,N \qquad (9)$$

The objective function is quadratic and the constraints are linear. This is a convex optimization problem. It can be solved using The Lagrange multiplier method. The objective function for this optimization problem (Lagrangian function) is:

$$Lp=\frac{1}{2}\|w\|^2 - \sum_{i=1}^{N}\lambda_i(y_i(w\cdot x_i+b)-1) \qquad (10)$$

Here, the parameters λ_i are called Lagrange multipliers.

We used SMO in our research as a SVM classifier.

A. SMO

SMO (Sequential Minimal Optimization) is an optimization algorithm used to train an SVM on a data set [12]. It was invented by John Platt in 1998 at Microsoft Research. SMO solves the SVM QP (quadratic programming) problems by decomposing it into QP sub-problems and solving the smallest possible optimization problem, involving two Lagrange multipliers, at each step [13]. These small QP problems are solved analytically, which avoids using a time-consuming numerical QP optimization as an inner loop [14]. SMO can handle a very large training set. This is allowed because the amount of memory required for SMO is linear in the training set size. Sequential Minimal Optimization (SMO) is a simple algorithm that can quickly solve the SVM QP problem without any extra matrix storage and without using numerical QP optimization steps at all [14]. SMO can fit very large SVM problems in the memory of personal computer or workstation. This algorithm is less susceptible to numerical precision problems because no matrix algorithms are used in SMO [14]. SMO has two components. It uses analytic method for solving the two Lagrange Multipliers. SMO also uses a heuristic for choosing which multipliers to optimize [14].

V. USED TECHNOLOGY AND DATASET

We used the Weka (Waikato Environment for Knowledge Analysis) tool in our research [15].

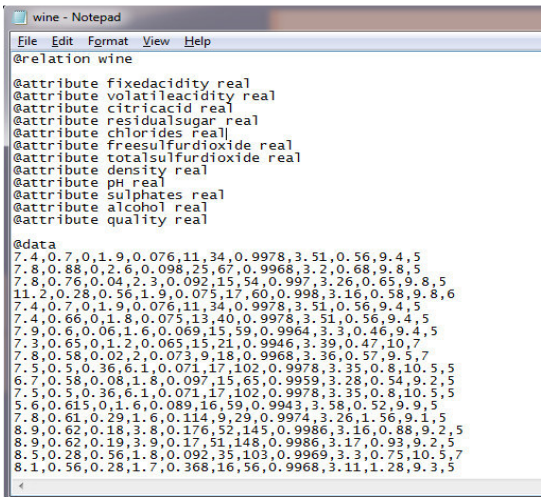

Figure 4. ARFF for wine quality dataset

We used the latest version 3.8.2. Weka contains implementation of machine learning algorithms for data mining tasks and a lot of data mining tools such as tools for preprocessing, classification, regression, clustering, association rules, and visualization. Weka is a software that is written in Java. It is open source software issued under the GNU General Public License. It also has a Graphical User Interface (GUI). This greatly helps in working with projects related to machine learning. For scripting large jobs, can be used Command Line Interface, which is part of Weka. We can integrate our applications with Weka using the Java API. Weka GUI has five parts: Explorer, Experimenter, Knowledge Flow, Workbench and Simple CLI.

TABLE II. DIFFERENCES BETWEEN THE CLASSIFICATION ALGORITHMS

Properties	ID3	SMO (SVM)
Total number of instances	1599	1599
Correctly Classified Instances	896	972
Correctly Classified Instances in percent	56.053%	60.788%
Incorrectly Classified Instances	483	627
Incorrectly Classified Instances in percent	39.212%	30.2064
Kappa statistic	0.4508	0.3583
Mean absolute error	0.1179	0.2346
Root mean squared error	0.3401	0.3294
Unclassified Instances	220	0

TABLE III. DETAILED ACCURACY BY CLASS FOR ID3 AND SMO

Detailed Accuracy By Class for ID3			
Precision	**Recall**	**F-Measure**	**Class**
0.000	0.000	0.000	3
0.102	0.125	0.112	4
0.724	0.722	0.723	5
0.666	0.655	0.661	6
0.556	0.556	0.556	7
0.167	0.182	0.174	8
Detailed Accuracy By Class for SMO			
Precision	**Recall**	**F-Meassure**	**Class**
0.000	0.000	0.000	3
0.250	0.019	0.035	4
0.657	0.764	0.706	5
0.571	0.603	0.587	6
0.532	0.332	0.409	7
0.000	0.000	0.000	8

For the purposes of our research, we used the wine quality dataset that is publicly available [16]. Specifically, in our research we classified red wine instances. Data refer to "Vinho Verde", a product that is produced in Minho, a district in northwestern Portugal. The dataset for red wine has 1599 instances. All instances have 12 attributes, of which 11 are physical-chemical properties of the wine, while the last feature is the wine quality. The wine quality can be in the range of 1-10 (1-very poor quality, 10-very good quality). In our research this feature is a class attribute. The other features are given in Figure 4.

Weka as a machine learning tool supports ".csv" and its natural ".arff" (Attribute-relation file format) types of input. An ARFF (Attribute-Relation File Format) file is an ASCII text file that describes a list of instances sharing a set of attributes [17]. In this research .ARFF was used as input. The structure of this file can be seen in Figure 4.

VI. OBTAINED RESULTS IN WEKA

For the purposes of this research, we used two classification algorithms: ID3 and SMO (SVM). As test option we used cross-validation with 10 folds. With the application of ID3, the total number of correctly classified instances is about 56%. The time it took to build the model is 0.03 seconds. Detailed accuracy results are given by Confusion Matrix:

```
=== Confusion Matrix ===

   a   b   c   d   e   f   <-- classified as
   0   2   3   1   0   0 |   a = 3
   4   5  20  11   0   0 |   b = 4
   3  30 433 118  16   0 |   c = 5
   0  12 118 361  55   5 |   d = 6
   0   0  24  47  95   5 |   e = 7
   0   0   0   4   5   2 |   f = 8
```

With the application of SMO (SVM), the total number of correctly classified instances is about 60%. The time it took to build the model is 2.78 seconds. Detailed accuracy results are given by Confusion Matrix:

```
=== Confusion Matrix ===

   a   b   c   d   e   f   <-- classified as
   0   1   7   2   0   0 |   a = 3
   1   1  31  19   1   0 |   b = 4
   0   1 520 147  11   2 |   c = 5
   0   1 213 385  39   0 |   d = 6
   0   0  20 111  66   2 |   e = 7
   0   0   1  10   7   0 |   f = 8
```

Using the Attribute Selection algorithm, we have obtained results by which the attribute "alcohol" has the greatest information gain. From Table 2 we can see the differences between ID3 and SMO algorithm. From Table 3 we can see the detailed accuracy by class.

VII. CONCLUSION

According to the obtained results SMO has more correctly classified instances than ID3, but ID3 has fewer incorrectly classified instances. The time of model building with SMO is greater than ID3. Using this training set, the quality of wine for new instances can be predicted in the future.

REFERENCES

[1] N. Jain , V. Srivastava, "Data mining techniques: a survey paper", IJRET: International Journal of Research in Engineering and Technology.

[2] B. Bringmann Mining patterns in structured data. https://lirias.kuleuven.be/bitstream/123456789/237427/1/CW2009_10.pdf

[3] I. Witten, E. Frank, Data mining : practical machine learning tools and techniques. Elsevier Inc, book.

[4] S. S. Nikam, "A Comparative Study of Classification Techniques in Data Mining Algorithms", Oriental journal of computer science & Technology, vol.8, pp.13-19, April 2015.

[5] S. Neelamegam, E.Ramaraj, "Classification algorithm in Data mining: An Overview", International Journal of P2P Network Trends and Technology (IJPTT), vol.4, pp.369-374, September 2013.

[6] J.Brownlee, "Supervised and Unsupervised Machine Learning Algorithms", https://machinelearningmastery.com/supervised-and-unsupervised-machine-learning-algorithms/

[7] F.Voznika, L.Viana, "Data mining classification", https://courses.cs.washington.edu/courses/csep521/07wi/prj/leonardo_fabricio.pdf

[8] K..Tekanomo, "Decision tree Tutorial / What is decision tree", http://people.revoledu.com/kardi/tutorial/DecisionTree/what-is-decision-tree.htm

[9] W.Peng, J.Chen, H.Zhou, "An Implementation of ID3 - Decision Tree Learning Algorithm", http://web.arch.usyd.edu.au/~wpeng/DecisionTree2.pdf

[10] The ID3 Algorithm, https://www.cise.ufl.edu/~ddd/cap6635/Fall-97/Short-papers/2.htm

[11] SVM (Support Vector Machine) – Theory. https://medium.com/machine-learning-101/chapter-2-svm-support-vector-machine-theory-f0812effc72

[12] SVM Tutorial, http://mccormickml.com/2013/04/16/trivial-svm-example/

[13] Fast Training of SVMs usin Sequential Minimal Optimization, http://www.d.umn.edu/~data0003/Talks/ml_svm.pdf

[14] J.C.Platt, "Sequential Minimal Optimization: A Fast Algorithm for Training Support Vector Machines", https://www.microsoft.com/en-us/research/wp-content/uploads/2016/02/tr-98-14.pdf

[15] Weka 3: Data Mining Software in Java, https://www.cs.waikato.ac.nz/ml/weka/

[16] Machine learning databases - Wine quality dataset, https://archive.ics.uci.edu/ml/machine-learning-databases/wine-quality/

[17] Attribute-Relation File Format (ARFF), https://www.cs.waikato.ac.nz/ml/weka/arff.html